

Prediction of bond dissociation energies and transition state barriers by a modified complete basis set model chemistry

Tim P. W. Jungkamp

Department of Environmental Engineering Science, California Institute of Technology, Pasadena, California 91125

John H. Seinfeld^{a)}

Department of Chemical Engineering, California Institute of Technology, Pasadena, California 91125

(Received 2 December 1996; accepted 25 April 1997)

The complete basis set model chemistries CBS-4 and CBS-q were modified using density functional theory for the geometry optimization step of these methods. The accuracy of predicted bond dissociation energies and transition state barrier heights was investigated based on geometry optimizations using the B3LYP functional with basis set sizes ranging from 3-21G(*d,p*) to 6-311G(*d,p*). Transition state barrier heights can be obtained at CBS-q with B3LYP/6-31G(*d,p*) geometries with rms error of 1.7 kcal/mol within a test set of ten transition state species. The method should be applicable to molecules with up to eight or more heavy atoms. Use of B3LYP/6-311G(*d,p*) for geometry optimizations leads to further improvement of CBS-q barrier heights with a rms error of 1.4 kcal/mol. For reference, the CBS-QCI/APNO model chemistry was evaluated and is shown to provide very reliable predictions of barrier heights (rms error=1.0 kcal/mol). © 1997 American Institute of Physics. [S0021-9606(97)02729-3]

I. INTRODUCTION

The fast developing field of *ab initio* methods in computational chemistry provides a number of techniques for the prediction of molecular properties. Depending on the size of the molecules being considered, the accuracy with which electronic energies are predicted can be very close to experimental values. However, with increasing number of heavy atoms, i.e., atoms heavier than hydrogen, less reliable predictions are obtained, since as a result of the exponential growth of the problem size with increasing number of electrons, only lower levels of theory with simplifying approximations can be applied.

Much effort has been directed toward the development of model chemistries, which allow the prediction of molecular properties using a uniform formalism for an entire size class of species. The first two models leading to chemical accuracy have been the G1 and G2 theories developed by Pople *et al.* (see Curtiss and Raghavachari for review).¹ The formalism common to G1 and G2 was developed to overcome uncertainties in electronic energy calculations as a result of the influence of basis set size and electron correlation. The key objective of these model chemistries is to extrapolate to a large scale calculation with relatively high electron correlation treatment and a large basis set, such as, e.g., QCISD(T)/6-311+G(3*df*,2*p*) in G2 theory. This calculation usually cannot be performed directly except for small problems, as in the G2 test set itself.² The extrapolation is obtained by calculating a series of single-point energies evaluating basis set size effects at a single level of theory (MP4SDTQ) and accounting for electron correlation treatment effects using a relatively small basis set

[6-311G(*d,p*)]. The combined trends of basis set influence and electron correlation treatment have been proven to extrapolate to the exact large scale calculation.² The resource-intensive G2 method has been revised several times in order to reduce basis set requirements (see Curtiss *et al.* and references therein).³

A family of modified G2 approaches, called G2M, was recently developed by Mebel *et al.*,⁴ replacing both the low level Hartree-Fock (HF) frequency calculation for determining the zero-point vibrational energies (ZPE) and the second-order Møller-Plesset (MP2) geometry optimization step in G2 by density functional theory (DFT) using the B3LYP functional. The QCISD(T) calculation is replaced by a spin restricted CCSD(T) calculation in G2M which further improves the accuracy of the method. This approach was also chosen by Bauschlicher and Partridge,⁵ leading to a similar formalism. G2M(RCC) is anticipated to perform well also in systems exhibiting serious spin contamination of the zeroth order wave functions, which is usual for transition states.

Independent of the development of the G2 methods, the approach of accounting for basis set influences by so called complete basis set extrapolation (CBS) was developed by Petersson *et al.*,⁶⁻⁸ with the aim of providing a standard technique with high predictive accuracy. Recently, the CBS method found general applicability in the form of the series of CBS model chemistries CBS-4, CBS-q, CBS-Q and CBS-QCI/APNO,⁶⁻⁸ each one adapted to a specific class of molecule size. The highest CBS model chemistry (CBS-QCI/APNO) provides the best absolute accuracy known today for molecules up to three heavy atoms with a mean average deviation (MAD) of 0.5 kcal/mol in atomization energies of 31 first row molecules.⁸ The accuracy of the G2 method (MAD=1.1 kcal/mol in atomization energies within the G2

^{a)} Author to whom correspondence should be addressed.

test set) can be achieved with less resources than needed for G2 by the CBS-Q method ($MAD=0.8$), most probably because a larger range of theory levels is covered in the CBS-Q extrapolation.

To be able to handle molecules of more than six heavy atoms, the CBS-4 model seems to be an appropriate choice, considering that it is capable of predicting bond dissociation energies of large molecules with essentially the same accuracy as G2.⁷ The HF/3-21G(*d*) geometry optimization underlying CBS-4 and CBS-q usually performs quite well due to a fortuitous cancellation of errors.⁹ However, this implies insufficient reliability in case of less well behaved molecules, like open shell species and especially transition states (TS).

The calculation of TS energies for sizable molecules remains a major challenge. The highly sophisticated electron correlation treatment provided by, e.g., multireference configuration interaction (MRCI), generalized valence bond (GVB), or complete active space self-consistent field (CASSCF) methods is not generally applicable. For these methods the computational problem scales to impracticable size with increasing number of basis functions, so that this type of treatment is limited to relatively small organic molecules. The results also may depend on, e.g., the choice of the active space of electron correlation in case of CASSCF calculations, which makes these methods difficult to apply as a general procedure belonging to a model chemistry.

It has become apparent that the application of model chemistries such as G2 is a useful tool for calculating transition state energies,^{1,10} at least for molecules of a size that can be treated with G2. Although the limit in number of heavy atoms is constantly expanding because of constantly decreasing hardware expenses and increasing CPU performance, still many important reaction mechanisms in organic chemistry cannot be addressed with existing methods. Most G2 derivatives wait to be proven reliable for transition states, with G2M(RCC) being the exception that was shown to reproduce the results of accepted high level results for some difficult problems with high spin contamination. However, the test set for G2M(RCC) comprised only four cases.⁴ To our knowledge, none of the G2(MP2) or related methods has been tested systematically for transition states.

The CBS family of model chemistries seems to be a potential candidate for transition state calculations, especially for larger molecules, but as yet no systematic check on their performance for transition states has been reported. The main focus of this paper is to modify the CBS-4 and CBS-q model chemistry to be able to predict TS energies by evaluating basis set requirements for the geometry optimizations. For testing the hypothesis, we applied modified CBS-q and CBS-4 model chemistries along with CBS-Q and CBS-QCI/APNO to a test set of relatively well known species and reactions.

The modification in the formalisms of CBS-q and CBS-4 is the use of DFT optimized geometries and zero-point vibrational frequencies (ZPE) instead of the HF geometries, which were part of the original formalism of these model chemistries. It is expected that the use of DFT for geometry optimization increases the reliability of these methods, since

correlation effects, which are clearly not treated by HF theory, are handled efficiently and economically by DFT methods. In addition, the efficiency of the DFT method should allow calculation of reliable transition state geometries for relatively large molecules. The comparison of CBS-q and CBS-4 should show whether the QCISD(T)/6-31G calculation necessary for CBS-q will significantly improve the CBS-4 energies. The latter model chemistry accounts for errors due to spin contamination only by an empirical correction measured by deviation of $\langle S \rangle^2$ of the HF/CbsB1 wave function. The CPU time for QCISD(T) calculations scales with the seventh power of number of atoms, whereas the MP4SDQ step used in CBS-4 scales only with the sixth power. Since the QCISD(T) component becomes the major fraction of CPU time needed for large molecules, it is important to know whether the QCISD(T) component is essential to achieve the desired level of reliability.

II. COMPUTATIONAL DETAILS

All calculations were performed using the parallel version of the GAUSSIAN94/LINDA program package.¹¹ DFT geometry optimizations and frequency calculations for the species of this study were performed using the Becke3LYP hybrid functional as implemented in GAUSSIAN94. Geometries of all species were optimized with tightened convergence cutoffs; transition states were optimized either using a redundant internal coordinate scheme or using eigenvalue following routines, both with analytical second-order gradients. ZPEs obtained at B3LYP/3-21G(*d,p*) and B3LYP/6-31G(*d,p*) were scaled with factors 0.9700 and 0.9828, respectively, derived from the G1 test of molecules by comparison with experimental ZPE values.¹² Unfortunately, the recent comprehensive study on scaling factors for vibrational frequencies and ZPEs by Scott and Radom¹³ did not include B3LYP/6-31G(*d,p*), but the result of 0.9828 is very close to the value of 0.9806 for B3LYP/6-31G(*d*), i.e., without polarization functions for hydrogen atoms. B3LYP/6-311G(*d,p*) ZPE values were taken without scaling.⁴ For CBS-QCI/APNO calculations the difference in spin contamination of the zeroth order wave functions of the CBS/(14s9p4d2f,6s3p1d)/[6s6p3d2f,4s2p1d] and QCISD(T)/6-311++G(2df,*p*) components have to be checked. In none of the cases considered did the $\Delta\langle S^2 \rangle$ of these two SCF wave functions exceed 0.005. Note that use of the APNO basis set for the QCISD(T) calculation, leading to a very expensive calculation, would be necessary for cases where $\Delta\langle S^2 \rangle > 0.01$. Care had to be taken when calculating the zeroth order wave functions at transition state geometries, since the likelihood of obtaining an unstable wave function in the SCF calculations was found to be very high for molecules with high symmetry and near degenerate electronic states. For this reason, stability checks were performed on a regular basis on all HF and B3LYP wave functions, using the techniques implemented in GAUSSIAN94.

The B3LYP hybrid functional B3LYP was used with 3-21G(*d,p*), 6-31G(*d,p*), and 6-311G(*d,p*) basis sets for geometry optimizations underlying the CBS model chemis-

TABLE I. Calculated and experimental 0 K atomization energies ΔE_a (kcal/mol) using the original and modified CBS model chemistries.

Molecule	CBS-4//		CBS-4//		CBS-q//		Exp ^a
	CBS-4	3-21	6-31	CBS-q	6-31		
O ₂	116.8	111.5	117.3	116.1	116.3	118.0	
CH ₃ CHO	644.2	644.0	644.7	642.2	642.7	642.6 ^b	
CH ₃ CO	556.0	556.0	556.3	553.8	554.1	558.1 ^b	
CH ₃ OH	481.1	480.8	481.4	480.3	480.5	480.8	
CH ₃ O	374.6	375.6	377.1	374.5	376.9	378.0 ^b	
CH ₂ OH	385.6	385.5	385.8	384.5	384.6	385.8 ^b	
H ₂ CO	358.0	357.9	358.6	357.1	357.7	357.2	
HCO	270.5	271.0	270.9	269.4	269.8	270.3	
C ₃ H ₈	945.8	946.9	946.2	943.9	944.3	942.9 ^b	
C ₃ H ₇	848.4	848.6	848.9	845.8	846.2	847.4 ^b	
C ₂ H ₆	668.0	667.9	668.2	667.3	667.6	666.3	
C ₂ H ₅	568.1	568.1	568.3	566.8	566.9	567.2 ^b	
C ₂ H ₄	532.4	532.9	533.1	531.2	532.1	531.9	
C ₂ H ₃	423.1	423.2	423.4	422.0	422.4	422.3 ^b	
C ₂ H ₂	385.6	386.2	387.0	384.5	386.2	388.9	
C ₂ H	255.9	255.6	255.3	254.0	255.3	255.6 ^b	
H ₂ O ₂	248.4	247.6	249.1	249.6	250.1	252.3	
H ₂ O	219.0	219.2	218.7	219.5	219.1	219.3	
OH	101.1	100.9	100.7	101.4	101.0	101.3	
CF ₄	466.1	463.9	466.2	461.4	461.6	465.5	
CH ₄	393.1	393.0	393.1	393.7	393.7	392.5	
CH ₃	289.5	289.1	289.2	289.4	289.1	289.2	
CH ₂ (¹ A ₁)	168.8	169.0	168.9	171.2	171.3	170.6	
CH ₂ (³ B ₁)	179.7	179.6	179.6	178.4	178.2	179.6	
CH	79.6	79.4	79.4	80.4	80.3	79.9	
NO	145.2	145.2	150.8	151.7	153.1	150.1	
NO ₂	222.7	221.8	221.8	227.0	226.1	221.9	
CN	172.1	171.3	171.6	177.2	176.6	176.6	
Average error	-0.5	-0.6	0.0	-0.6	-0.3		
MAD error	1.4	1.5	1.2	1.5	1.2		
rms error	2.0	2.2	1.6	2.1	1.7		
MAX error	4.9	6.5	3.3	5.1	4.2		

^aAs compiled in Ref. 8 unless otherwise noted.^bCalculated from ΔH_f (298 K) as compiled in Ref. 22, applying thermochemical corrections to the enthalpies obtained from B3LYP/6-31G(*d,p*) frequency calculations.

tries. This leads to modified CBS-4 and CBS-q formalisms, referred to as, e.g., CBS-4//3-21 in the case of CBS-4 using B3LYP/3-21G(*d,p*) geometry, or CBS-q//6-311 for geometry optimization, and ZPE determination from a frequency calculation at B3LYP/6-311G(*d,p*).

III. RESULTS AND DISCUSSION

A. Stable species

We have chosen a set of 28 small molecules containing H, C, N, O, and F for the benchmark calculations, reflecting our interest in oxidation pathways of hydrocarbons. Most of the species are members of the G2 test set, with additional species taken from the series of bond dissociation energy calculations of Ochterski *et al.*⁷

In Table I the atomization energies for the set of molecules calculated at CBS-4//3-21, CBS-4//6-31, and CBS-q//6-31 are compared with experimental values and those obtained with the original CBS-4 and CBS-q formalisms. For both CBS-4 and CBS-q models, the rms errors decrease by 0.4 kcal using the B3LYP/6-31(*d,p*) place of

HF/3-21G(*d*). This way, the reliability of the CBS energies benefits from the geometry optimizations using DFT. At a given choice of method for geometry optimization, the CBS-4 and CBS-q models perform quite similarly in terms of MAD and rms, e.g., for CBS-4/6-31 a rms error of 1.6 kcal/mol is found, which is essentially the same as 1.7 kcal/mol for CBS-q/6-31.

One might assume that improvement in geometries, and therefore CBS energies, is caused by the low extent of spin contamination for the DFT wave functions, regarding the considerable improvement found for NO in CBS-4/6-31 compared to CBS-4, since NO exhibits a high degree of spin contamination in the UHF wave functions at HF/3-21G geometry. However, it is clear from the results for other species like CN($\langle S^2 \rangle = 1.251$) and C₂H($\langle S^2 \rangle = 1.26$), that a high degree of spin contamination in the UHF/3-21G wave functions is not the primary cause for errors in geometries underlying the CBS energies. The major cause for deviations of CBS-4 energies can be found in the empirical measure of the effects of spin contamination on the convergence of the perturbation series expansion. Calculation of this spin contamination correction term $\Delta E = -38.43mE_h\Delta\langle S^2 \rangle$ is based on the deviation $\Delta\langle S^2 \rangle$ of the spin expectation value. The $\Delta\langle S^2 \rangle$ of the HF/CbsB1 wave function for NO at the HF/3-21G geometry ($r_{\text{NO}} = 1.2021 \text{ \AA}$) is 0.2775, compared to 0.0693 at B3LYP/6-31G(*d,p*) geometry ($r_{\text{NO}} = 1.1587 \text{ \AA}$). In contrast, $\Delta\langle S^2 \rangle$ for CN at HF/3-21G geometry ($r_{\text{CN}} = 1.1800$) is 0.5006, and is not significantly reduced at B3LYP/6-31G(*d,p*) geometry ($r_{\text{CN}} = 1.1739$) with $\Delta\langle S^2 \rangle = 0.4145$, so that at both HF and DFT geometry the CBS-4 energy remains spurious since $\Delta\langle S^2 \rangle$ is exceedingly high. Probably the assumption of a linear dependence upon $\Delta\langle S^2 \rangle$ in the spin correction term does not hold up to $\Delta\langle S^2 \rangle < 0.6$.⁸ Since spin contamination effects are effectively treated at the QCISD(T) level in CBS-q, the energies of those spin contaminated cases are inherently improved and close to experimental values. Also, the considerable reduction in error of atomization energy by 2.5 kcal/mol for CH₃O does not seem to be related to spin contamination since the UHF/CbsB1 wave functions are relatively pure doublet states at both HF/3-21G and B3LYP/6-31G(*d,p*) geometries ($\Delta\langle S^2 \rangle = .0095$ and .0100, respectively).

For calculating reaction energies, some of the errors apparent from the results for absolute energies are expected to cancel. Table II lists a total of 30 BDE values calculated for 0 K at the various levels of theory for our test set of C₁–C₃ hydrocarbons and oxygenated hydrocarbons. Also given in Table II, the variation of standard deviation in experimental errors can be taken into account calculating a weighted rms error as described in Ref. 7. The general enhancement found in rms errors of BDE values at DFT geometries represents a general trend, although it is mainly caused by a limited number of species, e.g., C₂H₄, C₂H₂, C₂H, CH₃O, and H₂O₂. With the present test set, the small deviations obtained even for the least resource demanding procedure, CBS-4//HF3-21G, are on the order of only 2.1 (weighted 1.3) kcal/mol rms from experimental BDE values and improve to 1.9 (weighted 0.9) kcal/mol using

TABLE II. Calculated and experimental 0 K bond dissociation energies BDE (kcal/mol) using the original and modified CBS model chemistries.

Molecule	Bond	CBS-4//		CBS-4//		CBS-q//		Exp (0 K) ^a
		CBS-4	3-21	6-31	CBS-q	6-31		
CH ₄	C-H	103.5	103.9	103.9	104.3	104.7		103.2±0.1
CH ₃	C-H	109.8	109.5	109.6	111.1	110.8		108.8±0.6
CH ₂	C-H	100.2	100.1	100.2	98.0	98.0		100.4±1.0
CH	C-H	79.6	79.4	79.4	80.4	80.3		79.9±0.1
C ₂ H ₆	C-H	99.8	99.8	100.0	100.6	100.7		99.8±0.5
C ₂ H ₅	C-H	35.7	35.2	35.1	35.5	34.8		34.4±0.5
C ₂ H ₄	C-H	109.3	109.7	109.8	109.3	109.8		109.7±0.8
C ₂ H ₃	C-H	37.6	36.9	36.3	37.6	36.2		33.6±0.8
C ₂ H ₂	C-H	129.6	130.7	129.9	130.4	130.9		131.5±0.7
C ₂ H	C-H	114.8	114.6	116.2	105.5	107.0		112.0±4.2
C ₂ H ₃ →CH ₂	C-H	80.0	79.6	79.5	78.4	77.9		81.0±3.5
CCH ₂ →C ₂ H	C-H	87.2	88.0	86.7	89.6	89.2		83.9±4.2
C ₃ H ₈	C ₂ -H	97.4	98.3	97.3	98.1	98.1		97.1±2.0
H ₂ CO	C-H	87.5	86.9	87.7	87.6	87.8		86.6±0.2
CH ₃ OH	C-H	95.5	95.3	95.5	95.9	95.9		95.9±0.3 ^b
CH ₃ COH	C-H	88.2	88.0	88.4	88.4	88.6		87.9±0.3
C ₂ H ₆	C-C	88.9	89.7	89.8	88.5	89.5		88.0±0.2
C ₂ H ₅	C-C	98.9	99.4	99.5	99.0	99.6		97.6±0.8
C ₂ H ₄	C-C	173.0	173.7	174.0	174.6	175.7		171.0±1.2
C ₂ H ₃	C-C	163.9	164.1	164.4	163.3	163.9		161.8±1.1
C ₂ H ₂	C-C	226.4	227.3	228.2	223.7	225.7		228.8±0.7
C ₂ H	C-C	176.4	176.1	177.7	173.6	175.0		177.4±0.8
C ₂	C-C	141.2	141.0	140.9	148.5	148.3		145.2±0.5
CCH ₂	C-C	163.4	164.0	164.2	165.3	166.2		160.7±4.2
CH ₃ CHO	C-C	84.1	83.9	84.6	83.3	83.8		82.7±1.2
CH ₃ OH	C-O	90.4	90.7	91.5	89.5	90.5		90.5±0.3
CO ₂	C-O	125.4	123.6	125.2	126.1	126.2		125.8±0.1 ^b
H ₂ O	O-H	117.9	118.2	118.0	118.1	118.1		118.1±0.1
H ₂ O ₂	O-H	91.1	89.9	86.9	87.3	85.2		86.5±0.2 ^b
CH ₃ OH	O-H	106.5	105.1	104.3	105.8	103.6		103.0±1.0
H ₂ O ₂	O-O	46.3	45.8	47.6	46.8	48.1		49.9±0.2 ^b
Average error		0.5	0.5	0.6	0.4	0.5		
MAD error		1.6	1.6	1.4	2.0	1.8		
rms error		2.1	2.1	1.9	2.7	2.4		
Weighted rms		1.3	1.2	0.9	1.2	1.0		
MAX error		4.6	4.2	4.3	6.5	5.5		

^aAs compiled in Refs. 6 and 7 unless otherwise noted.^bObtained from 298 K BDE with ΔH_f (298 K) values from Ref. 22 and thermochemical corrections at B3LYP/6-31G(*d,p*) level of theory.

B3LYP/6-31G(*d,p*) geometries. Again, no considerable difference between CBS-4 and CBS-q can be noticed, but B3LYP/6-31G(*d,p*) geometries improve the rms error of BDE by 0.3 kcal/mol. Surprisingly, cancellation of errors is more pronounced for the CBS-4 models, which consistently have a rms error lower by 0.5 kcal/mol compared to the CBS-q models. This promises that reactions of molecules up to eight heavy atoms could be investigated with high expected accuracy using contemporary computational resources.

B. Transition states

The test set for transition states included the relatively well known transition states for the following reactions: H+H₂, H+N₂, H+NO, NH+O, N+O₂, H atom transfer in HO₂, F+H₂, O+H₂, O+HCl, as well as OH and H elimination from HNNO. These transition states had been addressed by Durant and Rohlfing when evaluating the reliabil-

ity of G2 and its G2Q derivative.^{10,14} A subset of these species had been reinvestigated by Durant using DFT methods.¹⁵ We included the hydrocarbon radical hydrogen transfer reaction between C₂H₄ and C₂H₅ as an additional system that previously had been treated at a high level of theory. This system has several intrinsic difficulties imposed on a theoretical study, as there is severe spin contamination in the UHF wave functions and the triplet excitations are important.¹⁶

The highest level of theory applicable as a general procedure, CBS-QCI/APNO, was used in its original formalism to calculate transition state energies. The geometry optimizations were performed at the QCISD/6-311G(*d,p*) level, which is the same as that used by Durant and Rohlfing for their G2Q formulation of G2 theory (ZPE corrections for the TS were taken from the QCISD frequency calculations). The barrier heights found at the CBS-QCI/APNO level are compared to values from G2Q,¹³ G2M (where available),⁴ and

TABLE III. Barrier heights ΔE_0 for transition states obtained at different levels of theory using geometry optimizations as indicated along with literature values.

Transition state		Barrier height definition	CBS-4				CBS-q				CBS-Q	CBS-QCI/APNO	G2M (RCC) ^e	G2Q ^f	Previous work	
			HF/3-21G (d)	B3LYP/3-21G (d,p)	B3LYP/6-31G (d,p)	B3LYP/6-311G (d,p)	HF/3-21G (d)	B3LYP/3-21G (d,p)	B3LYP/6-31G (d,p)	B3LYP/6-311G (d,p)	MP2(FC)/6-31G	QCISD/6-311G (d,p)	B3LYP/6-311G (d,p)	QCISD/6-311G (d,p)		Ref.
H-H ₂	(² Σ)	TS-H ₂ -H	11.1	11.1	11.0	10.9	10.6	10.6	10.6	10.3	9.7	9.1		10.8	9.5	23
H-N ₂	(² A')	TS-N ₂ -H	12.2	11.2	12.1	12.5	12.0	11.3	12.3	10.9	14.2	14.5		14.5	15.2	24
H-NO	(³ A'')	TS-NO-H	5.6	5.6	0.5	4.1	4.5	0.0	1.6	4.9	1.9	2.8		4.1	4.1	25
NH-O	(³ Π)	TS-NH-O	-6.5	<i>N/G</i> ^a	-7.1 ^b	-8.8	8.5	<i>N/G</i> ^a	8.5 ^b	7.6	4.6	13.8		11.9	11.7	26
NH-O	(³ Π)	TS-NH-O	1.5	2.5	2.5	<i>NG</i>	0.4	2.2	4.6	<i>NG</i>	0.3	3.9		5.4	5.5	26
N-O ₂	(² A')	TS-O ₂ -N	-4.7	22.1	17.5	-16.5	9.9	11.9	9.8	6.9	3.0	11.2	6.3	6.2	11.7; 6.6	27;10
O-H ₂	(³ Π)	TS-H ₂ -O	9.6	8.2	11.4	12.3	10.0	8.9	11.5	12.1	11.6	10.5	15.5	14.8	12.7	28
O-HCl	(³ A'')	TS-HCl-O	9.3	9.4	12.0	12.6	8.6	8.6	10.1	10.4	12.7	<i>N/A</i> ^c		10.4	8.5	29
NN-OH	(² A')	TS-N ₂ O-H	<i>N/G</i> ^a	25.4	19.6	19.9	<i>N/G</i> ^a	18.4	15.2	14.8	12.9	17.0		16.6 ^g	16.3	17
H-NNO	(² A')	TS-N ₂ O-H	6.3	5.8	8.1	8.3	5.1	4.3	7.9	7.9	7.0	10.3		9.2 ^g	10.3	17
NN-OH	(² A')	TS-NH-NO	<i>N/G</i> ^a	-13.0	-14.5	-14.7	<i>N/G</i> ^a	-17.6	-19.7	-20.7	-22.4	-19.5		-17.9 ^g	-16.7	17
H-NNO	(² A')	TS-NH-NO	-29.4	-32.7	-26.1	-26.3	-27.2	-31.6	-27.0	-27.6	-28.3	-26.1		-25.3 ^g	-21.7	17
H-OO	(² A'')	TS-H-O ₂	12.3	9.5	11.5	11.1	8.5	5.5	8.3	8.9	7.6	8.7	3.4	1.8	9.2	30
O-H-O	(² A ₂)	TS-H-O ₂	-11.1	-14.5	-10.3	-10.1	-12.7	-16.0	-12.2	-12.1	-11.8	-12.3	-15.7	-17.0	-12.7	31
C ₂ H ₄ -H-C ₂ H ₄	(² A _g)	TS-C ₂ H ₄ -C ₂ H ₅	25.4	26.2	28.9	26.2	29.7	30.2	30.0	29.9	26.1	<i>N/A</i> ^d			27.7-30.7	16

^aNo saddle point located at this level of theory.^b³A'' electronic state with $\angle(\text{N-H-O}) = 172.2^\circ$.^cNot available, since APNO basis set only defined for first row atoms.^dNot available due to problem size.^eReference 4.^fReference 10.^gReference 14.

previous work, mainly obtained at the CASSCF/CCI level of theory (Table III). It can be seen that errors are usually quite small, with noticeable, but expected, deviations for the nonisogyric reactions exhibiting a high degree of spin contamination in the UHF wave functions of the TS. Prominent examples are the N-O_2 ($^2\text{A}'$) and NH-O ($^3\Pi$) TS. In the case of N-O_2 , it remains undecided how close the prediction at the CBS-QCI/APNO level is, since the extrapolated 0 K experimental activation energy (6.6 kcal/mol)¹⁰ and CASSCF/CCI barrier height (11.2 kcal/mol)¹⁷ differ by 5.1 kcal/mol. The CBS-QCI/APNO calculation seems to support the theoretical value, while the G2Q and DFT values by Durant^{10,15} reproduce the experimental value.

The CBS-QCI/APNO result for H-NNO seems to be closer to the G2Q and all other calculated values within this work than to the CASSCF/CCI value reported by Walch,¹⁷ which might be an insufficiency in the reference selection in the CCI procedure as discussed by Durant and Rohlfing,^{10,14} leading to an error of 3.0 kcal in the energy of the NH+NO asymptote. When referring the TS energy to the $\text{H+N}_2\text{O}$, the CBS-QCI/APNO and G2Q results are in excellent agreement with the CASSCF value.

We also applied the original formulation of the CBS-Q model chemistry to calculate the transition state barrier heights. Besides the troublesome cases noted above, the agreement is generally found to be acceptable, although the overall deviations increase considerably to an rms error of 2.4 kcal/mol, as compared to CBS-QCI/APNO with a rms error of 1.0 kcal/mol, which is probably caused by the MP2(FC)/6-31G⁺ geometry optimization. The choice of polarization functions taken from the 6-311G(*d,p*) basis set seems to have a deteriorating effect on the TS geometries, measured by the noticeable increase of rms bond length deviation compared to the MP2(FC)/6-31G(*d*) geometries (Table IV).

When reducing the level of theory for barrier height calculations in order to be able to treat molecules with more heavy atoms than possible with CBS-QCI/APNO, CBS-Q, or G2Q, care has to be taken on the choice of methodology. The calculation of transition state geometries is usually strongly dependent on both electron correlation treatment and basis set size. Since the HF/3-21G(*d*) geometry optimization used in the original CBS-4 and CBS-q method is *a priori* not expected to be capable of reproducing transition state geometries obtained at higher levels, density functional theory could result in a major improvement. Surprisingly, the HF/3-21G geometries are found to reproduce relatively close high level results, although the rms error in bond length is high compared to other methods used in this work. At the B3LYP/3-21G(*d,p*) level the location of the transition states is even less reliable than plain HF theory, and fails to locate proper transition state geometries in cases with loose TS. Geometries calculated at the B3LYP level using 6-31G(*d,p*) or 6-311G(*d,p*) basis sets are more reliable (Table IV), as can be seen by reduction in bond length rms errors by 30% and 40%, respectively, over HF/3-21G(*d*) (see Table IV).

It should be noted that density functional theory leads to

unusually long bond lengths compared to classic *ab initio* methods in the case of flat potentials.^{18–20} Clearly, the uncertainty in geometry has a major impact when used in hybrid methods such as the CBS model chemistries. Any hybrid model relies on the assumption that the geometries underlying the high level single-point calculations are close to the real minima, i.e., that the calculated geometry parameters are stable when applying higher correlation treatment levels. This is not the case for loose transition states, where the B3LYP and QCISD geometries differ by 100% or more. Also, at large bond distances the spin contamination in HF becomes much more pronounced than in DFT wave functions,²¹ leading to spurious results in the perturbation expansion for the single-point calculations that are based on loose DFT geometries. In cases of the tight transition states, the bond lengths found are much closer to high level values.

With respect to the barrier heights at all CBS levels considered in this work (Table III), the relatively small dependence of calculated barrier heights upon basis set size for geometry is somewhat unexpected. Only minor changes from B3LYP/6-31G(*d,p*) to 6-311G(*d,p*) occur due to the relatively close geometries. The description of the barrier heights involving highly spin contaminated UHF wave functions, i.e., N-O_2 ($^2\text{A}'$), and NH-O ($^3\Pi$), seem to be extremely inconsistent among the CBS models. Especially striking is the underprediction of barrier heights by the CBS-4 models, independent of geometry for those spin contaminated species. This is caused (as mentioned earlier for the stable but spin contaminated species) by the rough estimate of the effect of spin contamination upon the convergence of the perturbation series expansion in the CBS-4 model.

Without consideration of those difficult cases, however, there is a pronounced trend of improving accuracy, measured in terms of MAD, rms, and maximum error, with larger basis sets for geometry optimization and use of CBS-q instead of CBS-4, as shown in Table V. While the number of 13 species in the test set is still relatively small, the calculation of the MAD and rms from comparison of actually 10 TS, omitting the severely spin contaminated cases, is not meant to represent a real measure of statistic variability, rather it allows us to compare the individual methods we used here. From Table V it is apparent that use of CBS-q//6-31 seems to be a reasonable choice for tradeoff in CPU time needed for the level of geometry optimization required to achieve meaningful barrier heights. The rms for barrier heights at CBS-q is reduced from 4.1 kcal/mol by roughly 60% to 1.7 kcal using B3LYP/6-31G(*d,p*) geometries instead of HF/3-21G(*d*), whereas the reliability of the CBS-4 model does not improve using DFT geometries.

Some of the deviations found for the barrier heights are caused by systematic errors in the CBS energies of reference species. The H-N_2 transition state barrier heights are constantly underpredicted at all CBS-4 and CBS-q, regardless of geometry optimization procedure, since the energy of N_2 is overpredicted by 4.5 to 4.9 kcal/mol. Likewise, the large errors for the CBS-4 energy for NN-OH and H-NNO at B3LYP/3-21G(*d,p*) geometry are caused by the poor de-

TABLE IV. Geometries of transition states at different levels of theory. Bond lengths r given in Å, angles \angle in degrees.

Transition state	Parameter	HF/ 3-21G(d)	B3LYP/ 3-21G(d,p)	B3LYP/ 6-31G(d,p)	B3LYP/ 6-311G(d,p)	MP2/ 6-31G(d) ^a	MP2/ 6-31G†	QCISD/ 6-311G(d,p) ^a	Previous work ^a
H–H ₂ (² Σ)	$r(\text{HH})$	0.934	0.935	0.931	0.931	0.926	0.927	0.929	0.93
H–N ₂ (² A')	$r(\text{NN})$	1.132	1.129	1.123	1.115	1.117	1.116	1.131	1.15
	$r(\text{NH})$	1.420	1.602	1.550	1.519	1.337	1.329	1.418	1.398
	$\angle\text{HNN}$	113.5	119.7	119.0	119.3	117.3	122.7	117.3	122.7
H–NO (³ A'')	$r(\text{NO})$	1.182	1.195	1.161	1.152	1.161	1.153	1.165	1.16
	$r(\text{NH})$	1.753	4.270	2.262	2.153	1.582	1.570	1.738	1.85
	$\angle\text{HNO}$	109.1	121.6	116.9	117.7	125.0	124.6	116.0	114.0
NH–O (⁵ Π)	$r(\text{NH})$	1.255	1.067	1.056		1.145	1.156	1.131	1.11
	$r(\text{HO})$	1.325	1.811	1.936		1.336	1.317	1.361	1.446
NH–O (³ Π)	$r(\text{NH})$	1.171		1.123 ^b	1.126	1.099	1.107	1.094	1.139
	$r(\text{HO})$	1.415		1.345 ^b	1.333	1.466	1.441	1.467	1.36
N–O ₂ (² A')	$r(\text{OO})$	1.393	1.312	1.222	1.225	1.228	1.215	1.237	1.232
	$r(\text{NO})$	1.729	2.078	2.016	1.825	1.681	1.666	1.814	1.821
	$\angle\text{NOO}$	105.7	114.2	113.2	114.1	112.7	113.3	111.0	115.0
O–H–O (² A ₂)	$r(\text{OO})$	1.434	1.493	1.413	1.409	1.468		1.409	1.429
	$r(\text{H–OO})^c$	0.929	0.897	0.921	0.918	0.930		0.908	0.920
H–OO (² A'') ^d	$r(\text{OO})$	1.369	1.395	1.328	1.325	1.322	1.307	1.322	1.328
	$r(\text{H–OO})^c$	1.640	1.646	1.618	1.617	1.609	1.609	1.609	1.619
F–H ₂ (² Σ)	$r(\text{HH})$	0.940		0.744		0.799	0.811	0.788	0.765
	$r(\text{FH})$	1.128		2.332		1.323	1.286	1.380	1.539
O–H ₂ (³ Π)	$r(\text{HH})$	1.100	1.185	0.972	0.939	0.897	0.811	0.934	0.893
	$r(\text{HO})$	1.110	1.049	1.144	1.170	1.207	1.286	1.164	1.118
O–HCl (³ A'')	$r(\text{HCl})$	1.639	1.639	1.395	1.407	1.461	1.482	1.478	1.387
	$r(\text{HO})$	1.130	1.130	1.351	1.331	1.208	1.181	1.180	1.232
	$\angle\text{OHCl}$	180.0	180.0	136.4	139.6	139.7	140.8	137.7	133.4
NN–OH (² A')	$r(\text{NN})$		1.219	1.217	1.207	1.157	1.154	1.215	1.23
	$r(\text{NO})$		1.601	1.417	1.424	1.454	1.449	1.423	1.47
	$r(\text{NH})$		1.233	1.265	1.262	1.203	1.217	1.245	1.27
	$\angle\text{HNN}$		95.2	89.2	89.6	94.2	89.2	89.2	89.9
	$\angle\text{NNO}$		91.6	96.4	96.7	96.4	96.4	96.1	95.2
H–NNO (² A')	$r(\text{NN})$	1.170	1.156	1.148	1.141	1.129	1.127	1.152	1.18
	$r(\text{NO})$	1.232	1.234	1.195	1.186	1.197	1.187	1.191	1.21
	$r(\text{NH})$	1.641	1.778	1.662	1.647	1.423	1.422	1.547	1.47
	$\angle\text{HNN}$	108.9	115.1	114.1	114.2	116.8	116.9	111.3	109.6
	$\angle\text{NNO}$	161.9	175.4	172.6	172.0	168.6	167.8	167.4	160.2
C ₂ H ₄ –H–	$r(\text{CH})$	1.403	1.385	1.388	1.388	1.356 ^e	1.370		1.373 ^f
C ₂ H ₅ (² A)	$r(\text{CC})$	1.416	1.400	1.396	1.393	1.389 ^e	1.397		1.404 ^f
Bond length rms error ^g		0.101	0.137	0.070	0.060	0.039	0.051	0.041	

^aReference 10.^b³A'' electronic state with $\angle(\text{N–H–O}) = 172.2^\circ$.^cDistance between H and the center of OO bond.^dSecond-order saddle point.^eUMP2/6-31G(d,p) geometry, Ref. 16.^fUCCSD(T)/6-31G(d,p) geometry, Ref. 16.^gCalculated from bond length deviations from previous work without cases H–NO, NH–O, and N–O₂.

scription of N₂O at that geometry. The spurious CBS-q energy of O₂ at B3LYP/3-21G(d,p) geometry leads to the unusually high error observed for the H–OO and O–H–O TS.

IV. CONCLUSION

The enhanced accuracy of the CBS-4 and CBS-q model chemistries by using DFT optimized geometries is demonstrated by the decrease in MAD and rms for atomization energies and BDE values. It is mostly caused by the significantly better description for molecules like NO, C₂H₄, C₂H₂, C₂H, and CH₃O. The CBS-4 and CBS-q models were found to be capable of providing reliable estimates of barrier heights when DFT optimized geometries are used. The

smallest deviations within our test set were found for the CBS-q/6-31 model with an rms of 1.7 kcal in atomization energies, which allows the treatment of molecules of up to at least eight heavy atoms with moderate computational resources. In particular, the result for the C₂H₄–H–C₂H₅ TS is encouraging and indicates that reliable treatment of hydrocarbon reactions with CBS-q/6-31 should be possible. Use of B3LYP/6-311G(d,p) geometries reduces the rms error in CBS-q barrier heights to 1.4 kcal/mol. The use of CBS-q is recommended over CBS-4 since accuracy of barrier height prediction benefits from including the QCISD(T) contribution, as expected. However, where the QCISD(T) calculations become too large due to the problem size, CBS-4/6-31 still might be a reasonable approximation with an rms error

TABLE V. Deviation of barrier heights $\Delta(\Delta E_0)$ in kcal/mol from (and summary for) each level of theory.^a

Geometry	HF/ 3-21G(d)		B3LYP/ 3-21G(d,p)		B3LYP/ 6-31G(d,p)		B3LYP/ 6-311G(d,p)		MP2/ 6-31G†	QCISD/ 6-311G(d,p)	
	CBS- 4	CBS- q	CBS- 4	CBS- q	CBS- 4	CBS- q	CBS- 4	CBS- q	CBS-Q	G2Q	CBS- QCI/ APNO
H-H ₂	1.6	1.1	1.6	1.1	1.5	1.1	1.4	0.8	0.2	1.3	-0.4
H-N ₂	-3.0	-3.2	-4.0	-3.9	-3.1	-2.9	-2.7	-2.3	-1.0	-0.7	-0.7
H-NO	1.5	0.4	1.5	-4.1	-3.6	-2.5	0.0	0.8	-2.2	0.0	-1.3
O-H ₂	-3.1	-2.7	-4.5	-3.8	-1.3	-1.2	-0.4	-0.6	-1.1	2.1	-2.2
O-HCl	0.8	0.1	0.9	0.1	3.5	1.6	4.1	1.9	4.2	1.9	
NN-OH ^b			9.1	2.1	3.3	-1.1	3.6	-1.5	-3.4	-0.3	-0.7
H-NNO ^b	-4.0	-5.2	-4.5	-6.0	-2.2	-2.4	-2.0	-2.4	-3.3	-1.1	0.0
H-OO	3.1	-0.7	0.3	-3.7	2.3	-0.9	1.9	-0.3	-1.6	-7.4	-0.5
O-H-O	1.6	0.0	-1.8	-3.3	2.4	0.5	2.6	0.6	0.9	-4.3	0.4
C ₂ H ₄ -H-	-3.8	0.5	-3.0	1.0	-0.3	0.8	-3.0	0.7	-3.1		
C ₂ H ₅ ^c											
Average error	-0.6	-1.1	-0.4	-2.1	0.3	-0.7	0.6	-0.2	-1.0	-0.9	
MAD	2.5	3.1	3.1	2.9	2.4	1.5	2.2	1.2	2.0	2.1	0.8
rms error	2.7	4.1	4.0	3.4	2.6	1.7	2.5	1.4	2.4	3.1	1.0
Max error	-4.0	5.2	-4.5	-6.0	-3.6	-2.9	4.1	-2.4	4.2	-7.4	-2.2

^aCalculated from values given in Table III.^bUsing barrier height relative to N₂O+H.^c $\Delta(\Delta E_0)$ calculated as deviation from mean of 27.7–30.7 kcal/mol range of barrier height given in Ref. 16.

of 2.6 kcal/mol. The CBS-Q model chemistry exhibits 2.4 kcal/mol rms error in TS barrier heights, but might perform better with an improved geometry step. This error is actually smaller than the G2Q rms error of 3.1 kcal/mol within our test set (the latter value suffers from the poor description of the O₂(³Σ_g) energy for referencing the H-OO and O-H-O TS). Our goal was to find a method that is generally applicable for relatively large molecules and we did not seek to improve the CBS-Q results, since this method already requires considerable computational resources. The CBS-QCI/APNO model chemistry was found to provide accuracy in TS barrier height prediction better than G2Q, and seems to be the preferred choice with rms error of 1.0 kcal/mol whenever applicable, i.e., currently for systems with up to four first row heavy atoms.

ACKNOWLEDGMENTS

This work was supported by the U.S. Environmental Protection Agency, Center on Airborne Organics, under agreement R-819714-01-0. T.P.W.J. gratefully acknowledges a Forschungsstipendium der Deutschen Forschungsgemeinschaft. We thank Mr. Kiran Shekar for performing GAUSSIAN94 calculations. Allocation of CPU time for the IBM SP/2 by the Center for Advanced Computing Research (CACR) at the California Institute of Technology is gratefully appreciated.

¹L. A. Curtiss and K. Raghavachari, in *Quantum Mechanical Electronic Structure Calculations with Chemical Accuracy*, edited by S. R. Langhoff (Kluwer, Dordrecht, 1995).

²L. A. Curtiss, J. E. Carpenter, K. Raghavachari, and J. A. Pople, *J. Chem. Phys.* **96**, 9030 (1992).

³L. A. Curtiss, P. C. Redfern, B. J. Smith, and L. Radom, *J. Chem. Phys.* **104**, 5148 (1996).

⁴A. M. Mebel, K. Morokuma, and M. C. Lin, *J. Chem. Phys.* **103**, 7414 (1995).

⁵C. W. Bauschlicher and H. Partridge, *J. Chem. Phys.* **103**, 1788 (1995).

⁶J. A. Montgomery, J. W. Ochterski, and G. A. Petersson, *J. Chem. Phys.* **101**, 5900 (1994).

⁷J. W. Ochterski, G. A. Petersson, and K. B. Wiberg, *J. Am. Chem. Soc.* **117**, 11,299 (1995).

⁸J. W. Ochterski, G. A. Petersson, and J. A. Montgomery, *J. Chem. Phys.* **104**, 2598 (1996).

⁹W. J. Hehre, L. Radom, P. von R. Schleyer, and J. A. Pople, *Ab Initio Molecular Orbital Theory* (Wiley, New York, 1985).

¹⁰J. L. Durant and C. M. Rohlfing, *J. Chem. Phys.* **98**, 8031 (1993).

¹¹M. J. Frisch, G. W. Trucks, H. B. Schlegel, P. M. W. Gill, B. G. Johnson, M. A. Robb, J. R. Cheeseman, T. Keith, G. A. Petersson, J. A. Montgomery, K. Raghavachari, M. A. Al-Laham, V. G. Zakrzewski, J. V. Ortiz, J. B. Foresman, J. Cioslowski, B. B. Stefanov, A. Nanayakkara, M. Challacombe, C. Y. Peng, P. Y. Ayala, W. Chen, M. W. Wong, J. L. Andres, E. S. Replogle, R. Gomperts, R. L. Martin, D. J. Fox, J. S. Binkley, D. J. Defrees, J. Baker, J. P. Stewart, M. Head-Gordon, C. Gonzalez, and J. A. Pople, *GAUSSIAN94*, Revision C 4, Gaussian, Inc., Pittsburgh, PA, 1995.

¹²R. S. Grev, C. L. Janssen, and H. F. Schaefer, *J. Chem. Phys.* **95**, 5128 (1991).

¹³A. P. Scott and L. Radom, *J. Phys. Chem.* **100**, 16,502 (1996).

¹⁴J. L. Durant, *J. Phys. Chem.* **98**, 518 (1994).

¹⁵J. L. Durant, *Chem. Phys. Lett.* **256**, 595 (1996).

¹⁶J. D. Watts, J. A. Franz, and R. J. Bartlett, *Chem. Phys. Lett.* **249**, 496 (1996).

¹⁷S. P. Walch, *J. Chem. Phys.* **98**, 1170 (1993).

¹⁸K. D. Dobbs and D. A. Dixon, *J. Phys. Chem.* **98**, 12,584 (1994).

¹⁹B. Jursic, *Chem. Phys. Lett.* **256**, 603 (1996).

²⁰B. Jursic, *J. Chem. Phys.* **104**, 4151 (1996).

²¹J. Baker, A. Scheiner, and J. Andzelm, *Chem. Phys. Lett.* **216**, 380 (1993).

²²W. B. DeMore, S. P. Sander, R. F. Hampson, D. M. Golden, M. J. Kurylo, A. R. Ravishankara, C. J. Howard, C. E. Kolb, and M. J. Molina, *Chemical Kinetics and Photochemical Data for Use in Stratospheric Modeling*, JPL Publication 92-20 (Jet Propulsion Laboratory, Pasadena, CA, 1992).

²³A. I. Boothroyd, W. J. Keogh, P. G. Martin, and M. R. Peterson, *J. Chem. Phys.* **95**, 4343 (1991).

²⁴S. P. Walch, *J. Chem. Phys.* **93**, 2384 (1990).

²⁵S. P. Walch and C. M. Rohlfing, *J. Chem. Phys.* **91**, 2939 (1989).

²⁶S. P. Walch, *J. Chem. Phys.* **93**, 8036 (1990).

- ²⁷ S. P. Walch and R. L. Jaffe, J. Chem. Phys. **86**, 6946, (1987).
²⁸ S. P. Walch, J. Phys. Chem. **86**, 5670 (1987).
²⁹ H. Koizumi, G. C. Schatz, and M. S. Gordon, J. Chem. Phys. **95**, 6421 (1991).
³⁰ S. P. Walch, C. M. Rohlfing, C. F. Melius, and C. W. Bauschlicher, Jr., J. Chem. Phys. **88**, 6273 (1988); S. P. Walch and C. M. Rohlfing, *ibid.* **91**, 2373 (1989).
³¹ S. P. Walch and R. J. Duchovic, J. Chem. Phys. **94**, 7068 (1991).